ED 359 210 TM 019 836

AUTHOR Yap, Kim O.

TITLE Integrating Assessment with Instruction in ABE/ESL Programs.

PUB DATE Apr 93
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; *Administrators; *Adult Basic Education; Classroom Research; Community Colleges; Community Organizations; Correctional Institutions; *Educational Assessment; *English (Second Language); Instructional Design; *Integrated Activities; Interviews; Portfolios (Background Materials); Program Evaluation; Questionnaires; Writing Evaluation
IDENTIFIERS *Performance Based Evaluation; Washington

ABSTRACT

A research project was conducted in Washington State in 1992 to develop an assessment system that could be used by adult basic education (ABE) and English-as-a-Second-Language (ESL) programs to satisfy accountability requirements and to provide assessment information useful for classroom instruction. In this preliminary work, 35 ABE/ESL programs administered by community and technical colleges, community-based organizations, and correctional facilities participated. Eight individuals, generally program administrators, participated in key player interviews. Focus group discussions were attended by 9 teachers, 11 administrators, and 35 students. Questionnaires were completed by 35 administrators, 58 ABE teachers, 35 ESL teachers, 20 volunteers, 40 ABE students, and 25 ESL students. The following three assessment models were identified for making assessment useful for instructional purposes: (1) writing assessment; (2) portfolio assessment; and (3) classroom assessment. Implementation of these three models is likely to enhance the efficiency of ABE/ESL instruction while addressing accountability. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Kim O. Yap
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED359210

Integrating Assessment With Instruction In ABE/ESL Programs

Kim O. Yap
Northwest Regional Educational Laboratory

A paper presented at the annual meeting of the American Educational Research Association, Atlanta, Georgia, April 12-16, 1993

TM019836

BEST COPY AVAILABLE

INTEGRATING ASSESSMENT WITH INSTRUCTION IN ABE/ESL PROGRAMS

INTRODUCTION

Educational assessment originated from a need to detect individual differences for the purpose of sorting people into groups and predicting their future performance. During World War II, for example, the military depended heavily on assessment to prevent the "feeble minded" from joining its ranks. Since then, gatekeeping has been the raison d'etre of the assessment enterprise. In the 1970s and 1980s, with the advent of large scale assessments, attention shifted slightly from gatekeeping to accountability. A great deal of interest was focused on finding out if educational innovations really "worked." School people were pressured by policymakers to justify expenditures on innovative programs and, indeed, the entire system of public education. These pressures have persisted, perhaps even intensified, in the 1990s.

Thus, for decades, little emphasis has been placed on the use of assessment information for instructional improvement. Tyler, for example, lamented in 1951 that educational measurement has been a process "quite apart from instruction" instead of being an integral part of it (Tyler, 1951, p.47). More than 30 years later, Glaser again pointed out that with the exception of informal assessment conducted by good classroom teachers, assessments are typically not designed "to guide the specifics of instruction" (Glaser, 1986, p. 45).

More recently, a renewed interest in the use of assessment by classroom teachers to improve instruction has emerged (Nitko, 1989). This movement envisions assessment as an integral part of instruction. The focus is no longer on who will succeed or the reporting of results to meet accountability requirements. It is on helping the individual students gain the most from instruction (Linn, 1989).

Both the top-down accountability and the bottom-up instructional perspectives share a common goal of improving education. Accountability has been well established and has had a major impact on education. Using assessment information for instructional improvement, on the other hand, is a relatively new phenomenon and largely a research and development agenda -- a hope for the future (Linn, 1989).

THE RESEARCH PROJECT

This paper describes a research project conducted in 1992 in Washington to develop an assessment system that can be used by adult basic education (ABE) and English as a second language (ESL) programs to satisfy accountability requirements and to provide assessment information useful for classroom instruction. The project consisted of a review of relevant literature on adult basic education and assessment, and field research which included (a) interviews with key players in the assessment issue, (b) focus group discussions with educators and students in ABE/ESL programs, and (c) questionnaire surveys with ABE/ESL educators and students.

A total of 35 ABE/ESL programs administered by community and technical colleges, community-based organizations, and correctional facilities throughout the state of Washington participated in the field research. Eight individuals participated in the key player interviews. These were generally program administrators who were particularly concerned with assessment issues. The focus group discussions were attended by 9 teachers, 11 program administrators, and 35 students. Questionnaire surveys were conducted with 35 program administrators, 58 ABE teachers, 33 ESL teacher, 20 volunteers, 40 ABE students, and 25 ESL students. The teachers, volunteers and students were randomly selected for survey administration. The overall response rate for the surveys was over 85 percent.

FINDINGS

Literature Review

While assessment models abound in the literature which we have looked at (Davis & Yap, 1992; Davis & Arter, 1992), few appear to have the capabilities for promoting an integration between assessment and instruction. Of the few which hold promise in this regard, the following are particularly noteworthy:

Writing assessment

Recent advances in writing assessment have provided a viable vehicle for integrating assessment with instruction. Through a group process of reviewing student writings, the traits or targets of good writing have been identified and clarified for different grade levels. To train educators to use this trait model, sample papers are used to exemplify the different trait ratings. The assessment is linked to instruction and student outcomes through learning activities for each trait. Students internalize the traits by assessing their own and other students' papers.

Portfolio assessment

Portfolios, a relatively new approach, provide another means of integrating assessment and instruction. A consortium of educators, working under the auspices of the Northwest Evaluation Association, has defined a portfolio as:

> A (student) portfolio is a purposeful collection of student work that exhibits to the student (and/or others) the student's efforts, progress or achievement in (a) given area(s). This collection must include:
>
> - student participation in selection of portfolio content;
> - the criteria for selection;
> - the criteria for judging merit; and
> - evidence of student self-reflection.

Portfolios are assessment devices which clearly define targets/criteria of learning and use these criteria to portray student progress and achievement. Assembling portfolios encourages critical thinking and problem solving. Guidelines on what should be included in a portfolio have clear implications for teaching and learning.

Classroom assessment

A more general approach to integrating assessment and instruction, the classroom assessment model uses multiple strategies and measures to assess learners' progress. In addition to paper-pencil tests, alternative assessments--informal and formal--are used. Targets for learning activities are clearly defined and included in the instructional process.

Field Research

Key Players

The key players indicate that currently there is no link between assessment and instruction. For example, one of the most commonly used test, the Comprehensive Adult Student Assessment System (CASAS), does not match the ABE/ESL curriculum as reflected in the core competencies. Some key players feel that the core competencies are more focused on academic skills while CASAS items relate more to life skills.

A second concern is the lack of coordination among state agencies serving the adult student population. Currently, there is a lack of consistency in assessment approaches as well as in curriculum and instruction. Many key players expect Washington's Integrated Curriculum for Achieving Necessary Skills (ICANS) project to play a key role in addressing this concern. They view the new curriculum as a bridge between the core competencies (i.e., ABE competencies) and the Washington Basic Skills Competency System (i.e., JTPA competencies). The ICANS, along with the assessment work

conducted under the Northwest Workplace Basics project, is expected to provide a common curriculum for different state agencies to promote coordination and collaboration. Also, there is some expectation that the three key programs (ABE/ESL, JTPA, and JOBS) are moving toward competencies outlined in the Secretary's Commission on Achieving Necessary Skills (SCANS) report.

There is considerable agreement among the key players that the desired assessment model should have several critical features. First, it should serve as a means of improving instruction with a direct link between assessment and instruction. Second, it should empower students by providing them with feedback on the learning process and outcomes. A feedback loop to students is essential in this regard. Third, the model should include multiple measures to provide a comprehensive picture of student performance. Fourth, the model needs to be user-friendly. This includes ease of use and built-in staff development activities to ensure the proper use of assessment information. Fifth, the model should include benchmarks to depict student progress toward standards (e.g., attainment of certificate of initial mastery, meeting criteria to exit a learning level). Lastly, the model should provide consistency across projects and agencies.

Focus Groups

Focus group discussions with program administrators and teachers surfaced concerns very similar to those expressed by the key players. Specifically, these stakeholder groups are concerned about the absence of a link between assessment and instruction in the ABE/ESL programs. There is a great deal of interest in the use of assessment results for instructional improvement. Again, like the key players, these groups perceive a need for coordination among different state agencies serving the adult student population. Some expect the ICANS to provide the linkage.

Many favor the use of some form of standardized assessment approach, involving the use of multiple measures. Some believe that a high quality classroom assessment model will improve instruction by providing timely feedback to students. Staff training is seen as an essential component of any new assessment system.

While most of the participants are familiar with the traditional assessment approaches (i.e., standardized achievement tests), few are well versed in alternative assessment. Among current practices, the use of writing assessment, student folders, and some rudimentary form of portfolio assessment appear to hold promise for integrating assessment with instruction. Of those familiar with portfolio assessment, several mention student self-reflection as an essential ingredient.

The student participants indicate that a range of assessment devices are used for placement and diagnostic purposes. Among the most commonly used are the GED tests, ASSET, and interviews. Teacher-made tests are widely used in ESL programs.

It appears that very limited use is made of assessment information for instructional improvement. Verbal discussions of student progress appear to be a common practice. In writing assessment, holistic scoring is used occasionally. More often, writing samples are returned to students with comments only.

Many student participants indicate that their primary reason for taking the ABE/ESL classes is to upgrade job-related skills.

Questionnaire Survey

The survey of ABE/ESL program administrators, teachers, volunteers and students provided the following findings:

In most cases, structured interviews are conducted with students for placement purposes. In many cases, informal measures such as teacher-made tests and writing samples are also used. Formal measures used for placement purposes include TABE, CASAS, ASSET, BEST and ABLE. For ESL populations, the use of the Adult Refugee Test and teacher-made assessment devices is especially prevalent.

For student appraisal, a range of assessment measures are used. Formal measures include TABE, CASAS, ABLE, GED practice tests, and such norm-referenced tests as WRAT, PIAT, and Gates-MacGinitie. Informal measures include writing samples, curriculum-embedded tests, teacher-made tests, and teacher observation. Again, among ESL programs, the Adult Refugee Test and teacher-made tests are frequently used.

Classroom assessment involves a range of devices, including paper-pencil tests, worksheets, competency checklists, scoring of student products, student folders, oral questioning, quizzes, self-evaluation, observation of student performance, scoring class activities, and peer evaluation.

To assess mastery of competencies, the programs use formal and informal measures. Formal measures include TABE, CASAS, ASSET, ABLE, the Adult Refugee Test, and GED practice tests. Informal devices include writing sample, portfolios, teacher-made tests, oral interviews, and teacher observation.

The diverse range of assessment tools currently in use makes it difficult, if not impossible, to attain any level of consistency in assessment across the ABE/ESL programs.

For assessment to be effective in promoting student learning, the survey respondents indicate that it must have (a) a high degree of congruency with the curriculum, (b) a feedback loop to provide meaningful information to students, and (c) the flexibility to use multiple measures. In addition, assessment must be ongoing and non-threatening, preferably involving the student in the assessment process.

There is a sense of ambivalence among the respondents with regard to program evaluation. On the one hand, many recognize the need to have a common measure to look at accountability across programs. On the other hand, given the diversity of program and client characteristics, there is a perceived need to attend to this diversity by letting programs develop their own assessment strategies to achieve accountability. Program staff are especially doubtful that a fair evaluation can be conducted for ESL programs.

To improve current assessment practices, there is a general consensus that the programs need to go beyond paper-pencil tests to include alternative assessment approaches. There is also general agreement that to successfully implement any new assessment system, support, guidance, and staff training must be provided across the state.

Program administrators, teachers and volunteers appear moderately satisfied (with ratings of 3 and 4 on a 5-point scale) with the usefulness of the assessment information provided by current practices for placing students and identifying their needs. However, the satisfaction comes generally from the use of informal, locally developed measures. Many teachers indicate that teacher-made devices are much better than standardized measures.

A predominant majority of both ABE and ESL students find the current assessment practices helpful. Again, this is largely due to the use of such informal measures as checking worksheets, quizzes, reading writing assignments, keeping student folders, and reviewing homework assignments.

Self ratings of program staff's knowledge of assessment are moderately high (around 3.5 on a 5-point scale) and the use of assessment information to modify instruction appears moderately frequent (ratings of 3 and 4 on a 5-point scale). A predominant majority of the program administrators and teachers would like to receive training in classroom assessment.

Current assessment practices do not include adequate measures of critical thinking, problem solving, and interpersonal skills.

Given the complex characteristics of the adult student population and their general anxiety about test taking, there is a low level of trust in standardized instruments, particularly norm-referenced tests.

PROPOSED ASSESSMENT MODELS

The results of the literature review and field research suggest that assessment models, to be useful, must meet the following criteria:

First, the models should ensure a reasonable match between assessment and instruction in the ABE/ESL programs, an essential condition for integrating assessment with instruction.

Second, the models should provide consistency in assessment across programs, making it possible to aggregate assessment information up to the state level.

Third, the models should provide for the use of multiple measures and include built-in feedback loops to provide meaningful information to students to promote learning.

In the sections to follow, we provide a description of assessment models which meet these criteria and a discussion of their respective strengths and limitations.

Instructional Improvement Models

Writing Assessment

Recent advances in direct measurement of writing competencies have identified the traits or attributes of good writing. The traits cover the areas of:

- Ideas and content
- Organization
- Voice
- Word choice
- Sentence fluency
- Conventions

Each trait can be rated on a measurement scale. Each rating can be defined by a general statement followed by specific indicators. For example, a "3" rating on a 5-point scale may be defined by the following statement and indicators:

The reader can readily follow what is being said, but the overall organization may sometimes be ineffective.

- The introduction and conclusion are recognizable, though not so well crafted as the reader might wish.
- Placement or relevance of some details leaves the reader occasionally confused.
- The paper sometimes moves along at a good pace but at other times bogs down in trivia or speeds along too rapidly.
- Transitions sometimes work well; at other times, connections between ideas seem unclear.
- Despite problems, the organization does not seriously get in the way of the main point or storyline.

Samples of student writing (i.e., anchor papers) are used to demonstrate the traits and the different ratings at each grade level.

Results of assessment can be used to aid instruction on each trait through learning activities, using a process approach to writing (i.e., prewriting, writing, responding, revising, editing, postwriting, and evaluating). Students receive instruction on how to evaluate their own writing through self-reflection and peer evaluation. By understanding the traits and indicators, learners have a clear picture of what good writing looks like. Assessment is ongoing throughout the learning process. Information is documented and used to plan instruction.

Major strengths of the writing assessment model include the following:

- Providing clear targets of what is to be learned
- Having clear purposes of assessment
- Directly assessing writing competencies
- Getting the learner involved in the assessment process
- Traits being applicable across different modes of writing (e.g., narrative, expository, persuasive, letter writing, memos)

Some potential limitations of this model are that (a) anchor papers for the different traits, ratings, and modes need to be developed for ABE/ESL learners, and (b) time and resources need to be provided for extensive staff training in writing assessment.

Portfolio Assessment

"Portfolios" has come to mean different things to different people. Over the past two years, working under the auspices of the Northwest Evaluation Association, a consortium of educators has developed the following definition:

> *A (student) portfolio is a purposeful collection of student work that exhibits to the student (and/or others) the student's efforts, progress or achievement in (a) given area(s). This collection must include:*
>
> - *student participation in selection of portfolio content;*
> - *the criteria for selection;*
> - *the criteria for judging merit; and*
> - *evidence of student self-reflection.*

What a portfolio contains depends on the purpose of assessment. For example, the purpose of a portfolio assessment may be to evaluate a student for exiting a level, admission to college, or employment. Given this purpose, the student's best works would be included in the portfolio. If the portfolio is used to assess the processes a student goes through in developing a product, then drafts, revisions, and final products would be included.

Student participation in deciding what goes into the portfolio is essential because the instructional value of portfolios lies in the student's ability to apply criteria and to reflect on what has been learned.

Criteria or guidelines for what to include in the portfolio may be more or less structured. For example, each student may be required to include a letter to his/her landlord. Or the portfolio may include broad categories of work (e.g., letters, essays, poems) with students selecting a work sample for each category.

The decision to include or exclude a piece of work is based on some sort of criteria. Identifying and clarifying these criteria provides a process for the teacher and student to articulate what is valued in strong performance. With shared criteria, students become involved in an assessment process which empowers them to identify strong and weak performance and to improve their own performance.

In selecting samples of their work for a portfolio, students acquire skills in analysis, comparison, synthesis, problem-solving, and critical thinking. Documentation of self-reflection or oral reports on the process serve to motivate student learning.

Among its major strengths, portfolio assessment has the potential to:

- capture a wide array of complex student outcomes,
- display the processes which the student uses to produce work,
- enable the student to monitor his/her own progress and become part of the assessment process,
- align assessments with what is considered valued outcomes,
- provide ongoing and continuous information on student progress, and
- provide feedback to the teacher for instructional planning and improvement.

Some potential limitations or problems are (a) staff training is required to successfully implement the model, (b) a portfolio system needs to be carefully developed lest it become a mere collection of student work samples, and (c) teachers need to be sure that what is included in portfolios portrays a representative picture of student performance.

Classroom Assessment Model

A great proportion (more than 90 percent) of educational assessment is conducted by those who provide instruction to students -- the classroom teachers. Our field work with ABE/ESL programs indicates that the bulk of assessment activities is conducted with teacher-made devices (e.g., oral interviews, observation, oral questioning, quizzes). Yet, it has been well documented that most teachers have been poorly trained to conduct high quality assessment.

The classroom assessment model, a result of more than 10 years of research and development conducted by the Northwest Regional Educational Laboratory, seeks to alleviate the problem by designing and providing training to classroom teachers on how to conduct high quality assessment of student learning. While we have worked largely with the K-12 public school population, we believe the model applies equally well with the adult learner population. The model is based on sound assessment practices which research at NWREL and elsewhere has verified, including:

1. Clear targets or student outcomes. Valued achievement targets need to be explicitly defined. These targets may include knowledge to be mastered, thinking processes to be exhibited, behaviors to be demonstrated, products with specific attributes, or affective characteristics. Through a group process, teachers identify and clarify valued achievement targets. Curriculum and instruction are then developed to help students attain these learning targets and congruent assessment devices are constructed to assess their attainment.

2. Congruent assessment strategies. Assessment devices must be congruent with both the contents of the curriculum and the processes of instruction. For example, paper-pencil tests are efficient strategies for assessing factual recall and other discrete skills. Higher levels of learning must be assessed through alternative methods such as performance assessment. For example, an authentic measure of writing competencies can only be obtained through direct writing assessment. Such higher order skills as critical thinking and problem solving may need to be assessed through carefully designed student projects and oral interviews. Interpersonal skills may be assessed through interviews and observation of group interactions.

3. Student involvement. Students play a key role in the assessment process and their input is used in designing and implementing the assessment devices. Self evaluation and reflection on the learning outcomes stimulate and empower student learning as they provide feedback to the teacher for instructional planning and improvement.

4. Integration of assessment with instruction. The classroom assessment model uses assessment information on an ongoing, continuous basis for instructional planning and monitoring student progress.

As an inservice training model, classroom assessment offers a process that is applicable to most educational settings, including ABE/ESL programs. It is sufficiently comprehensive and flexible to address a wide range of student outcomes and to be workable with different teacher groups. Currently there are workshop packages and video presentations on a range of subject areas, including reading, mathematics, language arts, and science.

As a staff training package designed to upgrade assessment capabilities, the model is somewhat open-ended. Its effectiveness relies heavily on a sustained effort and a strong commitment on the part of program staff to improving the quality of assessment practices. Teachers need to be persuaded that it is worth their time and effort to learn how to conduct high quality assessment and to use assessment information for instructional planning and improvement.

Accountability Model

For accountability, we propose the use of a common set of criteria to derive a common metric aggregatable up to the state level. These criteria include:

- Validity of assessment procedures
- Reliability of assessment information
- Representativeness of assessment results
- Minimum error in data aggregation

A primary requirement with respect to validity of assessment procedures is that there be a reasonable match between testing and instruction. The contents of the assessment devices must be congruent with the ABE/ESL curriculum. Second, a common metric is necessary if data on program impact are to be aggregated to the state level to provide a statewide picture. Several metrics have been created for this purpose, including normative scores such as scale scores, percentiles, and normal curve equivalents and percentage of students scoring at particular levels of a measurement scale. Third, validity considerations require a procedure for identifying achievement gains attributable to the ABE/ESL programs. Since the use of a control group is impractical, if not impossible, a pretest-posttest design is a reasonable alternative. For example, a pretest measure can be taken at the time of intake or placement and a posttest measure taken after, say, 100 hours of instruction. The achievement gains measured in a common metric (e.g., scale scores) may then serve as an indicator of program impact. Other versions of the pretest-posttest design may also be feasible and meaningful.

Reliability of assessment information is assured when the instruments have documented evidence of reliability. Such evidence is typically presented in test manuals or reports on technical studies conducted on the assessment devices.

Assessment results are representative when the programs and students participating in the assessment constitute a sample representative of the target population with respect to such variables as type of program, geographic location, and program size. In many instances, a random sample of the target population will provide representative assessment results.

Data quality control measures (e.g., random checks on data recording, transcription, and aggregation) can be implemented to minimize errors which might be created in data collection and aggregation.

In the context of ABE/ESL assessment, the validity and reliability criteria can be applied in two ways. In the first option, a common measure (e.g., a single assessment device) may be identified on the basis of the criteria. For example, a single test may be identified among a set of candidate tests as the one which best meets the criteria.

In the second option, local programs may be given the flexibility of choosing their own test among a set of candidate tests. This option is desirable, perhaps even necessary in

some cases, given the diversity of program and student characteristics. An advantage of this option is that it makes it easier to maximize the congruency between testing and learning. However, any locally identified tests must meet the validity and reliability criteria.

In any case, accountability testing should be kept to a minimum. To this end, we recommend the use of a sampling or matrix sampling approach. In the first case (sampling), ABE/ESL programs can be sampled on a rotation basis. For example, a representative sample of one third of the programs may be selected on an annual or biennial basis for accountability assessment. Each program will be assessed for accountability once every three or six years. In the second case (matrix sampling), test burden on individual students is reduced by having each student take only a portion of the test, perhaps using procedures similar to those employed in the National Assessment of Educational Progress (NAEP). In either case, assessment information will be aggregated at program and state levels to provide a composite picture of program impact.

Although the accountability model is designed to provide summative information on program impact, the data can be rendered instructionally relevant through curriculum alignment (e.g., item analysis) to identify strong and weak areas of instruction at the appropriate level (e.g., program, class, student cohort, individual student). This approach would be particularly meaningful if the contents of the test are closely linked to curriculum and instruction in the ABE/ESL programs.

The implementation of the accountability model will initially arouse some suspicion and concern. Some of the concerns (e.g., lack of congruency between testing and teaching, student transiency) will be justifiable and need to be addressed in a positive and productive way. Others may stem from resistance to change and the fear that assessment data will be used to judge the effectiveness of program staff. These obstacles will need to be resolved by making clear to the program staff the purpose of program evaluation and the distinction between program accountability and personnel evaluation.

CONCLUDING REMARKS

The study reveals a rugged terrain of ABE/ESL assessment in the state of Washington. We suspect that many of the same concerns, issues, and problems also exist in other parts of the country. Through extensive field work with stakeholder groups and a comprehensive review of the extant literature, the study identifies three assessment models for making assessment useful for instructional purposes. The implementation of these models is likely to enhance the efficacy of ABE/ESL programs and at the same time address accountability issues where they really count -- what and how adult students are learning to become more productive citizens.

Across the nation, research and development efforts aimed at integrating assessment and instruction are still in an infancy stage (Nitko, 1989). This effort is no exception in many respects. What this paper presents are conceptual models which we think hold promise for ABE/ESL programs. We wish to make clear in presenting these models that much

work has yet to be accomplished before the concepts are translated into operational models.

References

Davis, A. E., & Yap, K. O. (1992). Results of Field Research on ABE/ESL Assessment. Portland, OR: Northwest Regional Educational Laboratory.

Davis, A. E., & Arter, J. A. (1992). Annotated Bibliography on ABE/ESL Assessment. Portland, OR: Northwest Regional Educational Laboratory.

Glaser, R. (1986). The integration of testing and instruction. *The redesign of testing for the 21st century: Proceedings of the 1985 ETS Invitational Conference*. Princeton, NJ: Educational Testing Service.

Linn, R. L. (1989). Current perspectives and future directions. In R. L. Linn (Ed.), *Educational measurement*. New York: American Council on Education/Macmillan.

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement*. New York: American Council on Education/Macmillan.

Tyler, R. W. (1951). The functions of measurement in improving instruction. In E. F. Lindquist (Ed.), *Educational measurement*. Washington, D.C.: American Council on Education.